




## RESEARCH ARTICLE

# Mitogenomes of Giant-Skipper Butterflies reveal an ancient split between deep and shallow root feeders [version 1; referees: 2 approved]

Jing Zhang<sup>1</sup>, Qian Cong<sup>1</sup>, Xiao-Ling Fan<sup>2</sup>, Rongjiang Wang<sup>3</sup>, Min Wang<sup>2</sup>,  
Nick V. Grishin  1,4

<sup>1</sup>Departments of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX, 75390-8816, USA

<sup>2</sup>Department of Entomology, South China Agricultural University, Guangzhou, Guangdong, 510640, China

<sup>3</sup>College of Life Sciences, Peking University, Beijing, 100871, China

<sup>4</sup>Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX, 75390-9050, USA

**v1** First published: 06 Mar 2017, 6:222 (doi: [10.12688/f1000research.10970.1](https://doi.org/10.12688/f1000research.10970.1))  
Latest published: 06 Mar 2017, 6:222 (doi: [10.12688/f1000research.10970.1](https://doi.org/10.12688/f1000research.10970.1))

## Abstract

**Background:** Giant-Skipper butterflies from the genus *Megathymus* are North American endemics. These large and thick-bodied Skippers resemble moths and are unique in their life cycles. Grub-like at the later stages of development, caterpillars of these species feed and live inside yucca roots. Adults do not feed and are mostly local, not straying far from the patches of yucca plants.



**Methods:** Pieces of muscle were dissected from the thorax of specimens and genomic DNA was extracted (also from the abdomen of a specimen collected nearly 60 years ago). Paired-end libraries were prepared and sequenced for 150bp from both ends. The mitogenomes were assembled from the reads followed by a manual gap-closing procedure and a phylogenetic tree was constructed using a maximum likelihood method from an alignment of the mitogenomes.

**Results:** We determined mitogenome sequences of nominal subspecies of all five known species of *Megathymus* and *Agathymus mariae* to confidently root the phylogenetic tree. Pairwise sequence identity indicates the high similarity, ranging from 88-96% among coding regions for 13 proteins, 22 tRNAs and 2 rRNA, with a gene order typical for mitogenomes of Lepidoptera. Phylogenetic analysis confirms that Giant-Skippers (Megathymini) originate within the subfamily Hesperiiinae and do not warrant a subfamily rank. Genus *Megathymus* is monophyletic and splits into two species groups. *M. streckeri* and *M. cofaqui* caterpillars feed deep in the main root system of yucca plants and deposit frass underground. *M. ursus*, *M. beulahae* and *M. yuccae* feed in the yucca caudex and roots near the ground, and deposit frass outside through a "tent" (a silk tube projecting from the center of yucca plant). *M. yuccae* and *M. beulahae* are sister species consistently with morphological similarities between them.

**Conclusions:** We constructed the first DNA-based phylogeny of the genus *Megathymus* from their mitogenomes. The phylogeny agrees with morphological considerations.

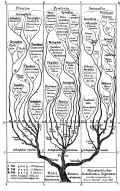
## Open Peer Review

Referee Status:  

	Invited Referees	
	1	2
version 1 published 06 Mar 2017	 report	 report
1	John A. Shuey, The Nature Conservancy USA	
2	Jia-Sheng Hao, Anhui Normal University China	

## Discuss this article

Comments (0)



This article is included in the **Phylogenetics** channel.

**Corresponding author:** Nick V. Grishin ([grishin@chop.swmed.edu](mailto:grishin@chop.swmed.edu))

**How to cite this article:** Zhang J, Cong Q, Fan XL *et al.* **Mitogenomes of Giant-Skipper Butterflies reveal an ancient split between deep and shallow root feeders [version 1; referees: 2 approved]** *F1000Research* 2017, **6**:222 (doi: [10.12688/f1000research.10970.1](https://doi.org/10.12688/f1000research.10970.1))

**Copyright:** © 2017 Zhang J *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Grant information:** This work was supported by National Institutes of Health [GM094575] and the Welch Foundation [I-1505]. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Competing interests:** No competing interests were disclosed.

**First published:** 06 Mar 2017, **6**:222 (doi: [10.12688/f1000research.10970.1](https://doi.org/10.12688/f1000research.10970.1))

Giant-Skippers (Lepidoptera: Hesperidae: Megathymini) are large, fat-bodied butterflies endemic to the North American continent<sup>1-3</sup>. Their caterpillars adapted to feeding inside large roots and fleshy leaves of *Yucca* and *Agave* plants and relatives. Protected from many predators living within their nutrition-rich food sources, Megathymini are larger in size than most other skippers, and don't feed as adults. Genus *Megathymus* is characterized by root-feeding caterpillars, mostly in *Yucca* plants, that build a "tent" (a silk tube projecting above the ground) at least prior to pupation. Caterpillars of the genus *Agathymus* live inside *Agave* leaves and make a "trap-door" (a round, hardened disk of silk) to close the entrance to their leaf chamber before pupation.

To better understand the evolution and phylogeny of *Megathymus*, we sequenced complete mitogenomes of all five known species from the genus: *M. yuccae*, *M. beulahae*, *M. ursus*, *M. streckeri*, and *M. cofaqui* (<http://www.butterfliesofamerica.com/L/Hesperidae.htm>). For most species, nominotypical subspecies from or near the type localities were used (see Figure 1 for specimen data; collected under the permit #08-02Rev). *M. beulahae* specimen, male paratype, was from the National Museum of Natural History collection (Smithsonian Institution, Washington, DC, USA). To confidently root the *Megathymus* tree, we also sequenced a complete mitogenome of *Agathymus mariae* as an outgroup. Methods for genomic DNA extraction, library construction, next-generation sequencing, and computational procedures followed those we reported previously<sup>4-14</sup>. The sequences have been deposited in GenBank and received accessions KY630500–KY630505.

All specimens, but one, were collected in 2013 and pieces of their muscles cut out of the thorax were preserved in 100% ethanol to ensure best DNA quality. However, *M. beulahae* paratype specimen was collected in 1957<sup>15</sup> and stored pinned, spread and dry in a museum drawer for 60 years. DNA was extracted from its abdomen prior to genitalia dissection and produced good quality genomic reads resulting in a complete mitogenome assembly. Similarly to the results reported previously<sup>16</sup>, we see that dry insect collections are an invaluable source of specimens for DNA studies; DNA can be extracted from Lepidoptera without damaging specimens beyond standard genitalia dissection procedure; and good quality DNA sequences can be obtained from specimens collected many decades ago.

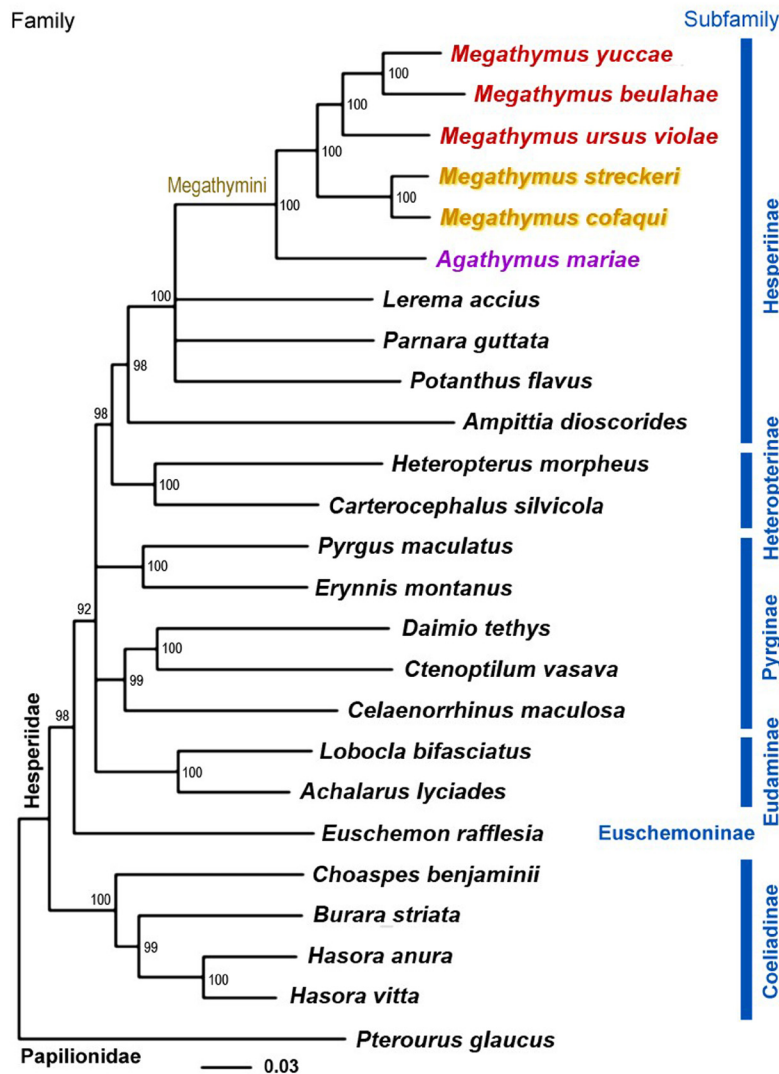
Sequence comparison revealed that mitogenomes of all six species of Megathymini were very similar, about 15K base pairs in length, coding for 13 proteins, 22 transfer RNAs and 2 ribosomal RNA with gene order typical for mitogenomes of Lepidoptera. The A+T-rich control region is most variable in sequence and length and contains several direct repeats of about 360 bp present in all six species. Among Hesperidae with available mitogenomes<sup>6,10,11,13,14,17-24</sup> these repeats are unique to Megathymini. The repeats cause difficulty with mitogenome assembly and their number remains uncertain.

To obtain the first DNA-based phylogeny of *Megathymus*, we constructed RAxML<sup>25</sup> (version 8.2.8, model GTRGAMMA, 100

bootstrap replicates) maximum likelihood tree from available high-quality mitogenomes of Hesperidae<sup>6,10,11,13,14,17-24</sup> rooted with *Pterourus glaucus* (Papilionidae) sequence<sup>10</sup> (Figure 1). While not giving confident resolution to the relationships between subfamilies Eudaminae and Pyrginae, the tree confirms the placement of Megathymini within the subfamily Hesperinae<sup>26-28</sup> and argues against historical treatment of Giant-Skippers at subfamily level. The tree resolves the Megathymini phylogeny with 100% bootstrap, supports monophyly of the genus *Megathymus* and suggests a split between the two species groups. The first group is formed by *M. streckeri* and *M. cofaqui*. Caterpillars of these species feed deep in the main root system of yucca plants and deposit frass underground<sup>2,3</sup>. They build a tent only prior to pupation and the tent usually projects from the ground surface. Males of these two species possess hair-like scales, particularly prominent on dorsal hindwing. *M. streckeri* and *M. cofaqui* are the closest sister species among *Megathymus* (Figure 1). Due to their apparently close relationship and allopatric distribution, Scott has suggested that *M. streckeri* and *M. cofaqui* may be subspecies of the same biological species<sup>2</sup>. However, the COI barcode sequences we obtained show about 4% divergence between them, revealing significant differences and supporting the two taxa as distinct species. COI barcode divergence in different populations of the same species mostly falls within 2%<sup>29</sup>.

The second species group consists of *M. yuccae*, *M. beulahae* and *M. ursus*. Caterpillars of these species feed in *Yucca* caudex and in roots close to the ground, maintaining the tent throughout development and depositing frass outside the tent<sup>2,3</sup>. Males of these three species lack hair-like scales. *M. yuccae* and *M. beulahae* are sister species, as expected from their close morphological similarities. However, their COI barcodes show a very large divergence of 9%. This pronounced divergence was unexpected, because the two species are quite similar in appearance and some males are difficult to tell apart ([http://www.butterfliesofamerica.com/L/t/Megathymus\\_a.htm](http://www.butterfliesofamerica.com/L/t/Megathymus_a.htm)). The most noticeable difference between *M. yuccae* and *M. beulahae* is the larger white ventral hindwing spots in the latter species, frequently fused to form a band, especially in females. However, these spots may be significantly reduced in males, frequently in the northern populations. Interestingly, *M. beulahae* is the only *Megathymus* species that feeds in yucca-like *Agave* plant<sup>1,15</sup>, but it is a confident sister of *Yucca*-feeding *M. yuccae*. *M. ursus* is a sister to these two *Megathymus*. *M. ursus* has rather different wing shape and patterns. The wings are narrower with more extended apex, forewing spots well-separated in *M. yuccae* form a band-like arrangement, and hindwings lack spots that females *M. yuccae* and *M. beulahae* possess.

In conclusion, we sequenced mitochondrial genomes of all five known species of *Megathymus* and one species of *Agathymus* as an outgroup, and constructed the first DNA-based phylogeny of *Megathymus*. The phylogeny is fully consistent with morphological and behavioral similarities between species. Our results support phylogenetic placement of Megathymini within the subfamily Hesperinae and clarify the relationships between *Megathymus* species. In particular, the major phylogenetic split is between the shallow



**Figure 1. Maximum likelihood tree of complete mitogenomes of 24 Hesperidae species rooted with *Pterourus glaucus* (Papilionidae).** Species names for mitogenome reported here are colored red. Numbers by the nodes show bootstrap support values and branches; bootstraps less than 70% are collapsed. GenBank accessions for sequences and data for specimens with mitogenomes reported here are: *Achalarus lyciades* NC\_030602.1; *Agathymus mariae mariae* KY630504, voucher NVG-1647, female, USA: New Mexico, Eddy County, 22-Sep-2013; *Ampittia dioscorides* KM102732.1; *Burara striata* KY524446; *Carterocephalus silvicola* NC\_024646.1; *Celaenorrhinus maculosa* NC\_022853.1; *Choaspes benjaminii* NC\_024647.1; *Ctenoptilum vasava* NC\_016704.1; *Daimio tethys* NC\_024648.1; *Euschemon rafflesia* KY513288; *Erynnis montanus* NC\_021427.1; *Hasora anura* NC\_027263.1; *Hasora vitta* NC\_027170.1; *Heteropterus morpheus* NC\_028506.1; *Lerema accius* NC\_029826.1; *Lobocla bifasciatus* NC\_024649.1; *Megathymus beulahae beulahae* KY630505, voucher 11-BOA-13385G05, paratype, male, Mexico, Hidalgo, near Ixmiquilpan, highway 85, klm. 176, 19-Aug-1957; *Megathymus cofaqui cofaqui* KY630503, voucher NVG-1536, female, USA: Georgia, Burke County, 2-Aug-2013; *Megathymus streckeri streckeri* KY630501, voucher NVG-1461, male, USA: Arizona, Apache County, southeast of Holbrook, 19-May-2013; *Megathymus ursus violae* KY630502, voucher NVG-1504, male, USA: Texas, Pecos County, Glass Mountains, 7-Jun-2013; *Megathymus yuccae yuccae* KY630500, voucher NVG-1185, male, USA: South Carolina, Aiken County, 25-Feb-2013; *Papilio glaucus* NC\_027252; *Parnara guttata* NC\_029136.1; *Potanthus flavus* NC\_024650.1; *Pyrgus maculatus* NC\_030192.1.

and deep yucca root feeders, and significant mitochondrial DNA divergences between *M. yuccae* and *M. beulahae* and between *M. streckeri* and *M. cofaqui* support the species status of these allopatric and similar in appearance taxa.

#### Author contributions

J.Z. assembled and analyzed the mitogenomes. Q.C. extracted DNA and prepared sequencing libraries. N.V.G conceived the project, and

drafted the manuscript. J.Z., Q.C., X-L.F., R.J.W., M.W. and N.V.G. analyzed the data and wrote the manuscript.

### Competing interests

No competing interests were disclosed.

### Grant information

This work was supported by National Institutes of Health [GM094575] and the Welch Foundation [I-1505].

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

### Acknowledgements

We acknowledge Texas Parks and Wildlife Department (Natural Resources Program Director David H. Riskind) for the permit #08-02Rev that makes research based on material collected in Texas State Parks possible. We are grateful to Robert K. Robbins, John M. Burns, and Brian Harris (National Museum of Natural History, Smithsonian Institution, Washington, DC) for granting access to the collections under their care. We thank Lisa N. Kinch for critical suggestions and proofreading of the manuscript. This work was supported in part by the National Institutes of Health (GM094575 to NVG) and the Welch Foundation (I-1505 to NVG).

### References

- Freeman HA: **Systematic review of the Megathymidae.** *J Lepid Soc.* 1969; **23**: 1–62.  
[Reference Source](#)
- Scott JA: **The Butterflies of North America: A Natural History and Field Guide.** (Stanford University Press); 1986.  
[Reference Source](#)
- Roeber K: **The Butterflies of North America.** (ed W. H. Howe); (Doubleday and Co.). 1975; 411–422.
- Cong Q, Borek D, Otwiniowski Z, *et al.*: **Tiger Swallowtail Genome Reveals Mechanisms for Speciation and Caterpillar Chemical Defense.** *Cell Rep.* 2015. pii: S2211-1247(15)00051-0.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cong Q, Borek D, Otwiniowski Z, *et al.*: **Skipper genome sheds light on unique phenotypic traits and phylogeny.** *BMC Genomics.* 2015; **16**: 639.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cong Q, Grishin NV: **The complete mitochondrial genome of *Lerema accius* and its phylogenetic implications.** *PeerJ.* 2016; **4**: e1546.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cong Q, Shen J, Borek D, *et al.*: **When COI barcodes deceive: complete genomes reveal introgression in hairstreaks.** *Proc Biol Sci.* 2017; **284**(1848): pii: 20161735.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cong Q, Shen J, Borek D, *et al.*: **Complete genomes of Hairstreak butterflies, their speciation, and nucleo-mitochondrial incongruence.** *Sci Rep.* 2016; **6**: 24863.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cong Q, Shen J, Warren AD, *et al.*: **Speciation in Cloudless Sulphurs Gleaned from Complete Genomes.** *Genome Biol Evol.* 2016; **8**(3): 915–931.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Shen J, Cong Q, Grishin NV: **The complete mitochondrial genome of *Papilio glaucus* and its phylogenetic implications.** *Meta Gene.* 2015; **5**: 68–83.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Shen J, Cong Q, Grishin NV: **The complete mitogenome of *Achalarus lyciades* (Lepidoptera: Hesperidae).** *Mitochondrial DNA B Resources.* 2016; **1**(1): 581–583.  
[Publisher Full Text](#)
- Shen J, Cong Q, Kinch LN, *et al.*: **Complete genome of *Pieris rapae*, a resilient alien, a cabbage pest, and a source of anti-cancer proteins [version 1; referees: 2 approved].** *F1000Res.* 2016; **5**: 2631.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Zhang J, Cong Q, Shen J, *et al.*: **The complete mitogenome of *Euschemon rafflesia* (Lepidoptera: Hesperidae).** *Mitochondrial DNA B Resources.* 2017; **2**(1): 136–138.  
[Publisher Full Text](#)
- Zhang J, Cong Q, Shen J, *et al.*: **The complete mitochondrial genome of a skipper *Burara striata* (Lepidoptera: Hesperidae).** *Mitochondrial DNA B Resources.* 2017; in press.
- Stallings DB, Turner JR: **A review of the Megathymidae of Mexico, with a synopsis of the classification of the family.** *The Lepidopterists' News.* 1958; **11**: 113–137.  
[Reference Source](#)
- Timmermans MJTN, Viberg C, Martin G, *et al.*: **Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections.** *Biol J Linn Soc.* 2016; **117**(1): 83–95.  
[Publisher Full Text](#)
- Cao L, Wang J, James John Y, *et al.*: **The complete mitochondrial genome of *Hasora vitta* (Butler, 1870) (Lepidoptera: Hesperidae).** *Mitochondrial DNA A DNA Mapp Seq Anal.* 2016; **27**(4): 3020–3021.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hao J, Sun Q, Zhao H, *et al.*: **The Complete Mitochondrial Genome of *Ctenoptilum vasava* (Lepidoptera: Hesperidae: Pyrginae) and Its Phylogenetic Implication.** *Comp Funct Genomics.* 2012; **2012**: 328049.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kim MI, Baek JY, Kim MJ, *et al.*: **Complete nucleotide sequence and organization of the mitogenome of the red-spotted apollo butterfly, *Parnassius bremeri* (Lepidoptera: Papilionidae) and comparison with other lepidopteran insects.** *Mol Cells.* 2009; **28**(4): 347–363.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kim MJ, Wang AR, Park JS, *et al.*: **Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperidae) and phylogenetic reconstruction of Lepidoptera.** *Gene.* 2014; **549**(1): 97–112.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Shao L, Sun Q, Hao J: **The complete mitochondrial genome of *Parara guttata* (Lepidoptera: Hesperidae).** *Mitochondrial DNA.* 2015; **26**(5): 724–725.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wang AR, Jeong HC, Han YS, *et al.*: **The complete mitochondrial genome of the mountainous duskywing, *Erynnis montanus* (Lepidoptera: Hesperidae): a new gene arrangement in Lepidoptera.** *Mitochondrial DNA.* 2014; **25**(2): 93–94.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wang J, James John Y, Xuan S, *et al.*: **The complete mitochondrial genome of the butterfly *Hasora anura* (Lepidoptera: Hesperidae).** *Mitochondrial DNA A DNA Mapp Seq Anal.* 2016; **27**(6): 4401–4402.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wang K, Hao J, Zhao H: **Characterization of complete mitochondrial genome of the skipper butterfly, *Celaenorrhinus maculosus* (Lepidoptera: Hesperidae).** *Mitochondrial DNA.* 2015; **26**(5): 690–1.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Stamatakis A: **RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.** *Bioinformatics.* 2006; **22**(21): 2688–2690.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Warren AD, Ogawa JR, Brower AVZ: **Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperidae (Lepidoptera: Hesperioidea).** *Cladistics.* 2008; **24**(5): 642–676.  
[Publisher Full Text](#)
- Warren AD, Ogawa JR, Brower AVZ: **Revised classification of the family Hesperidae (Lepidoptera: Hesperioidea) based on combined molecular and morphological data.** *Syst Entomol.* 2009; **34**(3): 467–523.  
[Publisher Full Text](#)
- Yuan X, Gao K, Yuan F, *et al.*: **Phylogenetic relationships of subfamilies in the family Hesperidae (Lepidoptera: Hesperioidea) from China.** *Sci Rep.* 2015; **5**: 11140.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Huemer P, Mutanen M, Sefc KM, *et al.*: **Testing DNA barcode performance in 1000 species of European lepidoptera: large geographic distances have small genetic impacts.** *PLoS One.* 2014; **9**(12): e115774.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

# Open Peer Review

Current Referee Status:  

---

## Version 1

Referee Report 27 March 2017

doi:[10.5256/f1000research.11826.r21272](https://doi.org/10.5256/f1000research.11826.r21272)



**Jia-Sheng Hao**

College of life sciences, Anhui Normal University, Wuhu, China

1. This work determined mitogenome sequences of nominal subspecies of all five known species of *Megathymus* and *Agathymus mariae* and constructed the phylogenetic trees of the main Hesperinae groups, the results confirms that Giant-Skippers (Megathymini) should not warrant a subfamily rank and that the monophyletic *Megathymus* splits into two species groups. In addition, the reconstructed mitogenomic phylogeny is fully consistent with morphological and behavioral similarities between the closely related species of the genus *Megathymus*. Overall this interesting work deserves to be published and indexed.
2. The manuscript's writing is good with an appropriate title and other content; the design, method and analysis is generally correct though some phylogenetic analyses should be more robust.
3. If the work considers the molecular dating of the *Megathymus* divergence by molecular clock method, and incorporates the analysis of relevant earth environmental factors, the results and significance will be more remarkable.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.

Referee Report 23 March 2017

doi:[10.5256/f1000research.11826.r21230](https://doi.org/10.5256/f1000research.11826.r21230)



**John A. Shuey**

The Nature Conservancy, Indianapolis, IN, USA

This paper is an interesting contribution to our understanding of the evolution of Megathymini and the phylogentic placement of the tribe within the Hesperidae. Using mitogenomes of all recognized species of *Megathymus*, the authors confirm morphologically-based species and species groups. Moreover – they identify a deep mitogenomic divergence between lineages that correspond with important life history traits that define species groups.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

***Competing Interests:*** No competing interests were disclosed.

---